## Cost-effectiveness of radiofrequency denervation for chronic low back pain

Maas, E.T.

2016

***document version***
Publisher's PDF, also known as Version of record

# 5

# Effectiveness and cost-effectiveness of radiofrequency denervation for chronic low back pain originating from the facet joints

Esther T Maas, Johan NS Juch, Raymond WJG Ostelo, J George Groeneweg,
Jan-Willem Kallewaard, Bart Koes, Arianne P Verhagen, Johanna M van Dongen,
Frank JPM Huygen, Maurits W van Tulder

Under review

## ABSTRACT

**Introduction** Radiofrequency (RF) denervation is widely used for treating low back pain (LBP) originating from the facet joints. The aim of this study is to establish whether RF denervation added to a standardised exercise programme is effective and cost-effective compared to a standardised exercise programme alone for patients with chronic low back pain (CLBP) originating from facet joints who are referred to a pain clinic.

**Methods** A multicentre, non-blinded randomised controlled trial and economic evaluation from a societal perspective was performed. Eligible patients were those who had not responded to conservative primary care and had CLBP originating from the facet joints (based on ≥50% pain reduction after a positive diagnostic block). Participants were randomly assigned to the intervention or control group, in which both groups received a standardised exercise programme of three months (8-12 hours) and psychological support if needed. The intervention group additionally received RF denervation. Primary outcomes were pain intensity (numeric rating scale 0-10), global perceived recovery (7-point Likert scale; 1-2 was defined as success), and functional status (Oswestry Disability Index) three months after the intervention. Participants completed questionnaires at baseline, three and six weeks, three, six, nine and 12 months after start of the treatment. Costs were collected using self-completed cost questionnaires. Longitudinal mixed-model analyses for the effects and seemingly unrelated regression analyses for the cost-utility were performed by intention to treat.

**Results** Between January 1, 2013 and June 3, 2014, 251 participants were randomised. Mean between-group differences adjusted for baseline characteristics were not statistically significant for pain intensity (at three months: -0.18; 95% CI -0.76 to 0.40), functional status (at three months: -2.45; 95% CI -5.93 to 1.03), and secondary outcomes at all measurements during the 12-month follow-up. For global perceived recovery, statistically significant effects favouring the intervention group were found at three weeks (OR 7.04; 95% CI 2.46 to 20.12) and six weeks (OR 3.28; 95% CI 1.43 to 7.52) after start of the treatment. After 12 months, there were no significant between-group differences in societal costs (€1184.51; 95%CI -78.10 to 2472.40) and no reasonable probability of cost-effectiveness was reached.

**Discussion** Both groups improved, however no between-group differences were found for pain and functional status at any follow-up measurement. Only statistically significant short-term (three weeks and six weeks) effects on global perceived recovery were found in favour of the intervention group. Our findings suggest that RF denervation is neither more effective nor cost-effective when added to a standardised exercise programme.

## INTRODUCTION

Globally low back pain causes more disability than any other condition, and has major social and economic consequences.[1-4] In The Netherlands the cost of low back pain is estimated at €3.5 billion in 2007, where the majority of the costs are due to patients developing chronic symptoms.[5] The lumbar facet joints are considered a possible cause of chronic low back pain (CLBP).[6,7] Facet joint pain is defined as pain resulting from any integral structure of the facet joints, including the fibrous capsule, synovial membrane, hyaline cartilage surfaces, and bony articulation.[8] Radiofrequency (RF) denervation is a common treatment for facet joint pain. RF denervation is a technique that attempts to modulate neural transmission of nociceptive stimuli, reducing spinal pain. It aims to denaturalise the nerves by applying an electric current (heat) to prevent the conduction of nociceptive impulses.[9,10]

Recent systematic reviews have shown low quality evidence for small positive effects of facet joint RF denervation on pain and functional status compared with placebo or steroid injections.[11-13] In clinical practice, RF denervation is typically provided in the setting of a multidisciplinary pain programme. Nonetheless, high quality evidence on the effectiveness and cost-effectiveness of facet joint RF denervation in addition to multidisciplinary pain programme is lacking.

The aim of this study is to establish whether RF denervation added to a standardised exercise programme is effective and cost-effective compared to a standardised exercise programme alone for patients with CLBP originating from facet joints who are referred to a pain clinic.



## METHODS

**Study design and participants**

This study is part of a larger collective initiative to evaluate minimal interventional treatments (MinT) in addition to a standardised exercise programme compared to a standardised exercise programme alone: the MinT study.[13] The MinT study consists of an observational study and three randomised controlled trials (RCTs) (1:1 ratio) with economic evaluations. Four RCTs were described in the published study protocol.[14] The study in which participants with disc problems were assessed was prematurely

terminated, as no participants with isolated disc problems were diagnosed after an inclusion period of five months. Ethical approval was granted by the Medical Ethics Committee of the Erasmus MC University Medical Centre in Rotterdam (registration number MEC-2012-079). All included participants gave written informed consent.

In 16 multidisciplinary pain clinics in the Netherlands, pain clinicians consecutively screened patients who sought care for LBP and were referred to a pain clinic. Inclusion criteria were: CLBP suspected to originate from the facet joints, age between 18 and 70 years, no improvement in symptoms after conservative treatment, and a positive (>50% pain reduction 30-90 minutes after procedure) diagnostic block.

*The diagnostic block*: A 22G needle was inserted to the posterior primary root of the spinal nerve (medial branch) under C-arm fluoroscopy. L3, L4, and L5 were selected for diagnostic blocks. The lateral image was checked to confirm the correct position of the needle, after which 0.5ml 2% lidocaine was injected. No corticosteroids were administered. The diagnostic block was considered positive if the patient reported at least 50% subjective pain reduction 30 minutes after the block, after which the patients were randomised.

Exclusion criteria were: pregnancy, inability to complete questionnaires, anticoagulant drug therapy and/or coagulopathy, body mass index (BMI) >35, involvement in a work-related conflict, and severe psychiatric or psychological problems. More details on the eligibility criteria can be found in the study protocol.[14]

**Randomisation and masking**

To detect a clinically relevant mean difference of two points on the NRS for pain intensity (SD 4), a total of 85 participants per group was needed (using a power of .9, alpha .05 and a correlation of .5 for repeated measurements).[15] Anticipating potential study withdrawal (20%), 102 participants per group were required. A centrally developed computerised random number generator was used for randomisation, and stratified for the pain clinics (N=16). Participants who met the inclusion criteria and gave informed consent were allocated by a local research nurse in each pain clinic to the intervention or control group (1:1 ratio).

In this pragmatic trial, participants and care providers were not blinded. All outcome measures were self-reported by the participant, therefore the outcome measurement was not blinded either. To ensure data were stored and analysed anonymously, all participants were assigned a unique number. Conclusions based on data analyses made

by the project team were blinded for treatment allocation. Patient expectations and satisfaction were measured to evaluate a possible risk of bias of having a non-blinded trial.

**Interventions**

All participants received the multidisciplinary pain programme consisting of a standardised exercise programme and psychological support if necessary (based on expert opinion). Intervention group participants also received the RF denervation. Anaesthesiologists at the participating pain clinics conducted the diagnostic blocks and the RF denervation. Every participating pain clinic had a referral agreement with physiotherapy practices in their region to provide the standardised exercise programme.

*Exercise programme:* All participants received a standardised exercise therapy programme based on the guideline on LBP of the Royal Dutch Society for Physiotherapy.[16] As the application of these guidelines can vary between physiotherapists, a standardised version was developed in cooperation with experienced physiotherapists. The entire duration of the programme ranged between 8 to 12 hours spread evenly over three months. More details can be found in the study protocol.[14]

*Psychological support:* If necessary, the participant was referred to a psychologist. The participant received psychological support to the discretion of the psychologist.

*Radiofrequency denervation:* Supplementary to the exercise programme, the intervention group received RF denervation, based on the current guideline on anaesthesiological pain management control in the Netherlands within one week after the first exercise therapy session.[17] A C-arm image intensifier was positioned in a slightly (10–15°) oblique position to identify skin entry points with the patient in prone position. A 22 G SMK needle with a 10-mm active curved tip was introduced at each entry point. The position of the cannula was checked on the lateral and AP fluoroscopic projection. The depth was adjusted until the tip of the cannula was at the level of a line connecting the posterior aspects of the intervertebral foramen. Sensory stimulation was positive if muscle contraction occurred below 0.6V. Second stimulation at 2 Hz was used in which contraction of the musculus multifidus and no leg contractions should occur. Once the position of the electrode was satisfactory, 1-2 ml per level ml 2% lidocaine was injected and a 90°C 90 seconds RF

lesion was made of the medial ramus dorsalis of L3-4, L4-5, and L5-S1.
In both treatment groups, participants were asked to refrain from any co-interventions during the three-month intervention period. Co-interventions after the initial intervention period were monitored and evaluated. If participants in the control group had not improved after three months, it was possible for them to return to the general practitioner or pain clinic and they could receive other treatment. All of this was registered.

**Outcomes**

Primary outcomes were scored on the 11-point Numerical Rating Scale (NRS) for pain intensity,[18] 7-point Likert scale for global perceived effect (GPE),[19] and the Oswestry Disability Index (ODI) for functional status,[20] which are included in the core outcome set for clinical trials in LBP.[21] Pain reduction of more than 30% or two points decrease in NRS was defined as treatment success for a post-hoc responder analysis. Success of treatment in GPE was achieved if a patient scored "complete recovery" or "much recovery".
Secondary outcomes were measured by the EuroQol (EQ-5D) for health-related quality of life,[22] 7-point Likert scale for patient satisfaction,[23] Rand-36 for general health,[24] and the Multidimensional Pain Inventory (MPI) for chronic pain experiences.[25] The participants' EQ-5D health states were converted into utilities using the Dutch tariff.[26] Quality Adjusted Life Years (QALYs) were calculated using linear interpolation between measurement points.
Intervention costs, other healthcare utilisation, informal care, unpaid productivity and absenteeism due to back pain were collected using three-monthly self-reported cost questionnaires to perform a cost evaluation.[27] Intervention costs were estimated using two hospital accounting records. Healthcare utilization included primary care, secondary care, and the use of prescribed and over-the-counter medication. Healthcare utilization was valued using Dutch standard costs and prices of professional organizations if unavailable.[28] Medication use was valued using unit prices of the Royal Dutch Society of Pharmacy.[29] Informal care included care by family, friends, and other volunteers and was valued using a recommended Dutch shadow price of €13.7/hour.[28] Absenteeism was measured with the Productivity and Disease Questionnaire (PRODISQ).[30] In accordance with the friction cost approach (friction period=23 weeks), absenteeism was valued using age- and gender-specific price weights.[28] Unpaid productivity costs included all hours of

volunteer work, domestic and educational activities that participants were not able to perform due to their CLBP and was valued using a recommended Dutch shadow price of €13.7/hour.[28] All costs were converted to 2014 Euros using consumer price indices.[31] All web-based questionnaires were sent at baseline, three, six, nine, and 12 months after the start of the treatment. Pain intensity, GPE and QALYs, were also assessed at three and six weeks after start of treatment.

**Statistical analyses**

*Effectiveness:* Analyses were by intention-to-treat and all available data were used. We adjusted for the possible effect of missing data on the study results in the analysis of mean changes for continuous outcomes by using maximum likelihood estimation for longitudinal mixed-effects models under "missing at random" assumptions and including a term for pain clinic if necessary.[32] Additionally, we performed a post-hoc responder analysis using a generalized linear mixed model (logit link) with the same multilevel structure. Regression coefficients or odds ratios with 95% confidence intervals (CI) were calculated. We adjusted for baseline characteristics and the effect of interest for this study was the time by treatment interaction. Data were compared between complete and incomplete cases to identify possible selective drop-out. Additionally, a sensitivity analysis estimated an "as-treated" longitudinal analysis based on comparisons of those actually treated, excluding the participants in the control group who did receive RF denervation. Effectiveness data were analysed in MLwiN (V2.22) with a level of significance of $p<.05$.

*Cost-effectiveness:* Cost-utility analyses (CUA) were performed by intention-to-treat, and from a societal perspective. Baseline characteristics were compared between intervention and control group participants.[33,34] Missing data were handled using multiple imputation. The imputation model included gender, smoking, marital status, age, BMI, complaint history, education, treatment expectations, and available baseline and effect measure values. Using Fully Conditional Specification and Predictive Mean Matching, 10 complete data sets were created (loss-of-efficiency<5%). Pooled estimates were calculated according to Rubin's rules.[35] Mean between-group cost differences were calculated for total and disaggregated costs. Seemingly unrelated regression analyses were performed in which effect and cost differences were corrected for baseline characteristics while taking into account the possible correlation between costs and effects.[36] Incremental

cost-utility ratios (ICERs) were calculated by dividing the corrected difference in total costs by the difference in QALYs. Uncertainty surrounding the cost differences and ICERs were estimated using Bias Corrected and Accelerated (BCA) bootstrapping techniques (5000 replications). The latter was graphically presented in cost-effectiveness planes.[37] Cost-effectiveness acceptability curves (CEACs) were estimated to indicate the interventions' probability of cost-effectiveness at different values of willingness-to-pay.[38] Sensitivity analyses were performed comparing the SF-6D and EQ-5D (SA1), comparing the friction cost approach to the human capital approach (SA2), and by performing a complete-case analysis (SA3), in order to assess the robustness of the results. The economic evaluation was performed using STATA (V12, Stata Corp, College Station, TX).

**Role of the funding source**

The funder of the study had no role in study design, data collection, data analysis, interpretation of data, or writing of the paper. The corresponding author had full access to all the data in the study and had final responsibility for the decision to submit for publication.

## RESULTS

**Study participants**

In total 7103 patients were assessed for eligibility; a total of 2013 patients were included in the MinT study (one of the three RCTs or observational study) between January 1, 2013 and June 3, 2014 (the inclusion period for patients in the RCTs assessing facet joint complaints). Of these 2013 participants, 865 participants fulfilled the criteria for receiving a diagnostic facet joint block. The block was positive in 630 participants, 251 of those participants met all inclusion criteria for this RCT assessing facet joint complaints and were randomly assigned to the intervention (N=125) and control group (N=126) (Figure 1). The patients with a negative diagnostic facet joint block continued the diagnostic trajectory or were followed-up in the observational study.

Baseline characteristics were comparable in both groups (Table 1). However, participants appeared to have LBP complaints longer in the intervention group. Length of back pain complaints was 146 months in the intervention group, and 100 months in the control group.

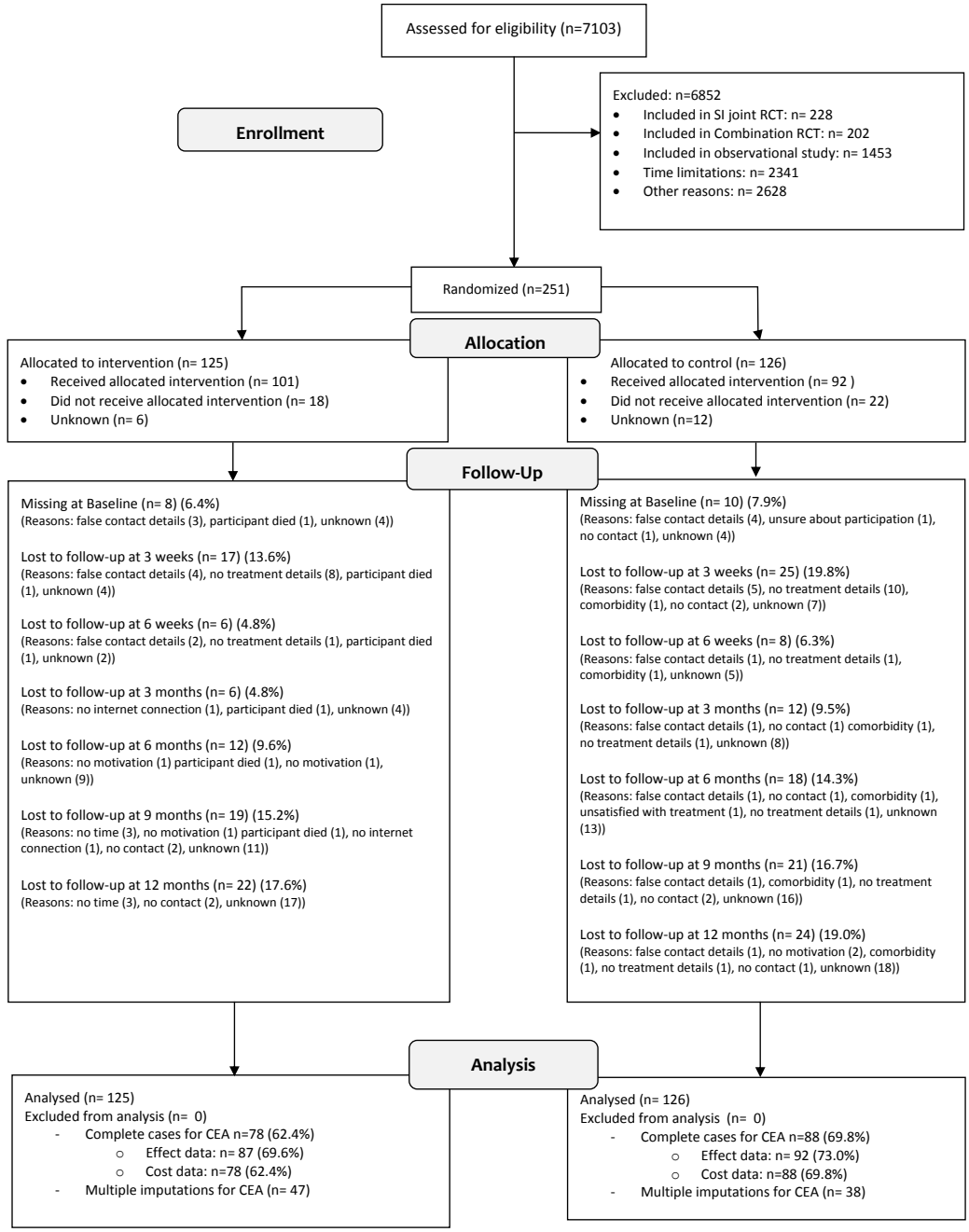



***Figure 1. Flow diagram***

Complete data on the primary outcomes after three months was obtained from 233 participants (93%). Complete data outcome on all follow-up moments during the year, was obtained from 179 participants (71%) (intervention n=87; control n=92) on the effect measures and from 166 (66%) on the cost measures (intervention n=78; control n=88). The participants with complete data were older, more often non-smokers, were more likely to have a partner, had a higher BMI, and had back pain complaints for a longer period of time (Table S1 in the Supplementary Appendix). These variables were included in the imputation model.

During the three-month intervention period, 12 participants in the control group received the RF denervation and were considered protocol violators. Between three- and 12-months follow–up, 31 control group participants (24.6%) received a RF denervation (26 participants received facet joint RF denervation, five participants received SI-joint RF denervation), and 32 intervention group participants (25.6%) received another RF denervation (30 participants received a second facet joint RF denervation, two participants received SI-joint RF denervation).

*Table 1. Baseline characteristics*

| Characteristics | Intervention; N=125* | Control; N=126* |
|---|---|---|
| **Age in years (SD)** | 52.9 (11.5) | 52.6 (10.8) |
| **Female (N (%))** | 65 (55.6%) | 60 (51.7%) |
| **BMI (SD)** | 26.7 (5.2) | 27.6 (4.3) |
| **Smoker (N (%))** | 34 (29.1%) | 34 (29.3%) |
| **Education** | | |
| • Low (N (%)) | 57 (48.7%) | 64 (55.2%) |
| • Moderate (N (%)) | 35 (29.9%) | 34 (29.3%) |
| **History of back pain complaints** | | |
| • Time since first experience with low back pain in months (median (IQR)) | 146 (50-267) | 100 (37-186) |
| • Time since current episode with low back pain in months (median (IQR)) | 31 (12-103) | 27 (11-73) |
| **Married/living with a partner (N (%))** | 93 (74.4%) | 98 (77.8%) |
| **Expectations** | | |
| • Credibility (0-27) | 21.4 (3.9) | 19.5 (5.5) |
| • Expectancy (0-27) | 18.9 (4.6) | 17.4 (5.2) |
| **Having a paid job** | 64 (51.2%) | 66 (58.9%) |
| **Outcomes** | | |
| **Pain intensity in the past week (NRS 0-10)** | 7.1 (1.4) | 7.2 (1.3) |
| **Oswestry disability index (mean (SD))** | 35.1 (14.7) | 34.4 (12.2) |
| **Quality of life (EQ-5D)** | 0.5 (0.3) | 0.5 (0.3) |

*Abbreviations: SD, Standard Deviation; N, number; BMI, Body Mass Index; EQ-5D, EuroQol-5D*
** Results are presented of the 233 participants who had complete baseline data*



### Intention-to-treat analyses

Pain intensity reduced in both groups: in the intervention group from 7.14 at baseline to 5.01 after three months and to 4.49 after 12 months. In the control group, pain intensity reduced from 7.19, to 5.44 and 4.44, for the three and 12 months follow-up, respectively. No statistically significant differences were shown between the two groups for pain intensity (NRS: 0-10) three months after the intervention (MD -0.18; 95%CI:-0.76 to 0.40), and all other follow-up assessments (Table 2). For success based on GPE, statistically significant results in favour of the intervention group were found three weeks (OR 7.04; 95%CI: 2.46 to 20.12) and six weeks (OR 3.28; 95%CI:1.43 to 7.52) after start of the treatment. For functional status, no statistically significant differences between the two groups were found at any follow-up moment. No statistically significant differences for all secondary outcomes at any follow-up moment were found (Table 4).

No adverse events or treatment complications were reported during the one year follow-up.

***Table 2. Treatment effects for primary outcomes based on intention-to-treat analyses***

| | | Mean Intervention group (SD) | Mean Control group (SD) | Treatment effect (95%CI) | P value for difference | |
|---|---|---|---|---|---|---|
| **Primary outcomes** | | | | | | |
| **NRS Pain*** | Overall effect | | | -0.08 (-0.50 to 0.34) | 0.71 | |
| | Baseline | 7.14 (1.38) | 7.19 (1.30) | | | |
| | 3 weeks | 5.17 (2.27) | 5.92 (1.72) | -0.41 (-1.02 to 0.19) | 0.18 | |
| | 6 weeks | 5.19 (2.31) | 5.90 (1.98) | -0.38 (-0.96 to 0.20) | 0.20 | |
| | 3 months | 5.01 (2.29) | 5.44 (2.21) | -0.18(-0.76 to 0.40) | 0.55 | |
| | 6 months | 4.61 (2.89) | 4.84 (2.42) | -0.04 (-0.63 to 0.56) | 0.91 | |
| | 9 months | 4.66 (2.41) | 4.73 (2.53) | 0.19 (-0.41 to 0.80) | 0.53 | |
| | 12 months | 4.49 (2.48) | 4.44 (2.56) | 0.47 (-0.14 to 1.07) | 0.13 | |
| **ODI Functioning*** | Overall effect | | | 0.04 (-3.02 to 3.10) | 0.98 | |
| | Baseline | 35.08 (14.66) | 34.40 (12.23) | | | |
| | 3 months | 26.03 (16.58) | 28.67 (15.25) | -2.45 (-5.93 to 1.03) | 0.17 | |
| | 6 months | 25.38 (15.64) | 27.15 (16.16) | -0.60 (-4.13 to 2.92) | 0.74 | |
| | 9 months | 25.74 (15.55) | 24.52 (15.56) | 2.26 (-1.29 to 5.82) | 0.21 | |
| | 12months | 24.59 (16.28) | 25.04 (16.64) | 1.48 (-2.09 to 5.06) | 0.42 | |
| | | **Success % Intervention group** | **Success % Control group** | **OR (95%CI)** | **P value for difference** | **NNT** |
| **GPE Success** | Overall effect | | | 1.39 (0.93 to 2.09) | 0.11 | |
| | 3 weeks | 29.6 | 5.0 | 7.04 (2.46 to 20.12) | 0.0003 | 4 |
| | 6 weeks | 29.4 | 9.3 | 3.28 (1.43 to 7.52) | 0.005 | 5 |
| | 3 months | 36.1 | 23.7 | 1.52 (0.76 to 3.05) | 0.24 | 8 |
| | 6 months | 40.7 | 36.1 | 1.07 (0.54 to 1.20) | 0.85 | 22 |
| | 9 months | 38.7 | 40.0 | 0.72 (0.36 to 0.44) | 0.35 | -77 |
| | 12months | 42.7 | 39.2 | 0.85 (0.43 to 1.70) | 0.65 | 29 |

*Values presented are model estimates of linear mixed-effects models with a random intercept, and adjusted for baseline and age, gender, BMI, education, smoking, marital status, back pain complaint history, patient expectations and baseline values. Regression coefficients can be interpreted as mean differences between interventions at a certain follow-up moment compared to baseline. Abbreviation: SD: Standard Deviation; NRS, Numeric Rating Scale; GPE, Global Perceived Effect; ODI, Oswestry Disability Index; NNT, Numbers Needed to Treat * Higher score indicates more severe symptoms. Range for NRS pain, 0-10; for ODI, 0-100; for GPE, 1&2 is success*

### Responder analyses

In terms of success, no differences between the groups were found for pain intensity (Table 3).

***Table 3 Treatment effects for primary outcomes based on intention-to-treat analyses, in terms of successful treatment***

| | | Success % Intervention group | Success % Control group | OR (95%CI) | P value for difference | NNT |
|---|---|---|---|---|---|---|
| **Primary outcomes** | | | | | | |
| **NRS Pain success, >30% reduction** | Overall effect | | | 1.10 (0.74 – 1.64) | 0.62 | |
| | Baseline | | | | | |
| | 3 weeks | 39.2 | 27.0 | 1.51 (0.75 – 3.07) | 0.25 | 9 |
| | 6 weeks | 40.2 | 31.6 | 1.20 (0.62 – 2.30) | 0.59 | 12 |
| | 3 months | 45.1 | 36.7 | 1.28 (0.67 – 2.43) | 0.46 | 12 |
| | 6 months | 55.6 | 50.5 | 1.05 (0.55 – 2.01) | 0.88 | 20 |
| | 9 months | 40.2 | 31.6 | 1.20 (0.61 – 2.37) | 0.60 | 12 |
| | 12 months | 47.0 | 53.5 | 0.62 (0.32 – 1.21) | 0.16 | -15 |
| **NRS Pain success, >2 points reduction** | Overall effect | | | 0.98 (0.67 – 1.44) | 0.93 | |
| | Baseline | | | | | |
| | 3 weeks | 54.9 | 44.0 | 1.36 (0.70 – 2.62) | 0.36 | 10 |
| | 6 weeks | 50.9 | 41.2 | 1.16 (0.62 – 2.16) | 0.65 | 11 |
| | 3 months | 56.1 | 46.8 | 1.14 (0.61 – 2.13) | 0.68 | 11 |
| | 6 months | 63.0 | 58.1 | 1.01 (0.53 – 1.92) | 0.98 | 21 |
| | 9 months | 54.9 | 56.9 | 0.79 (0.41 – 1.51) | 0.47 | -50 |
| | 12 months | 55.0 | 63.6 | 0.58 (0.30 – 1.13) | 0.11 | -12 |

*Abbreviation: NRS, Numeric Rating Scale; NNT, Numbers Needed to Treat*

### Sensitivity analyses

The analyses were replicated without the 12 protocol violators, however the results did not change on any primary outcome measure at any follow-up moment (Table S2 in the Supplementary Appendix). The analyses have been replicated without patients receiving the intervention after the three-month intervention period; this did not alter the results either (Table S3 in the Supplementary Appendix). Analysing only complete cases showed statistically significant between-group differences for pain intensity at 6 weeks (-0.77; 95%CI:-1.48 to -0.06) (Table S4 in the Supplementary Appendix).

***Table 4 Treatment effects for secondary outcomes based on intention-to-treat analyses, continuous***

| | | Mean Intervention group (SD) | Mean Control group (SD) | Treatment effect (95%CI) | P value for difference |
|---|---|---|---|---|---|
| **Secondary outcomes** | | | | | |
| **EQ5D Utilities*** ** | Overall effect | | | 0.01 (-0.03 to 0.04) | 0.75 |
| | Baseline | 0.52 (0.26) | 0.54 (0.26) | | |
| | 3 weeks | 0.69 (0.21) | 0.64 (0.23) | 0.05 (-0.01 to 0.10) | 0.08 |
| | 6 weeks | 0.69 (0.20) | 0.67 (0.21) | 0.03 (-0.03 to 0.08) | 0.32 |
| | 3 months | 0.68 (0.25) | 0.69 (0.21) | -0.01 (-0.06 to 0.05) | 0.85 |
| | 6 months | 0.73 (0.21) | 0.71 (0.23) | 0.02 (-0.03 to 0.07) | 0.42 |
| | 9 months | 0.72 (0.22) | 0.75 (0.18) | -0.05 (-0.10 to 0.01) | 0.11 |
| | 12 months | 0.73 (0.22) | 0.73 (0.22) | -0.03 (-0.08 to 0.03) | 0.37 |
| **Patient satisfaction**** | Overall effect | | | -0.01 (-0.30 to 0.28) | 0.96 |
| | Baseline | - | - | | |
| | 3 months | 2.95 (1.37) | 3.26 (1.41) | -0.18 (-0.54 to 0.18) | 0.34 |
| | 6 months | 2.96 (1.18) | 3.06 (1.35) | 0.01 (-0.35 to 0.38) | 0.94 |
| | 9 months | 2.88 (1.28) | 3.12 (1.49) | -0.02 (-0.39 to 0.35) | 0.91 |
| | 12 months | 2.88 (1.42) | 3.01 (1.43) | 0.19 (-0.19 to 0.56) | 0.32 |
| **MPI Pain severity**** | Overall effect | | | 0.05 (-0.21 to 0.32) | 0.70 |
| | Baseline | 89.69 (17.59) | 87.51 (17.41) | | |
| | 3 months | 80.25 (18.27) | 81.00 (19.76) | -0.15 (-0.48 to 0.18) | 0.36 |
| | 6 months | 77.83 (19.26) | 77.94 (19.98) | -0.03 (-0.37 to 0.31) | 0.86 |
| | 9 months | 77.30 (21.22) | 76.38 (21.25) | 0.16 (-0.18 to 0.50) | 0.36 |
| | 12 months | 76.09 (21.46) | 76.47 (21.67) | 0.27 (-0.07 to 0.61) | 0.12 |
| **MPI interference**** | Overall effect | | | -0.06 (-0.31 to 0.19) | 0.63 |
| | Baseline | 3.30 (1.18) | 3.12 (1.15) | | |
| | 3 months | 2.62 (1.32) | 2.74 (1.34) | -0.15 (-0.44 to 0.14) | 0.32 |
| | 6 months | 2.59 (1.39) | 2.57 (1.38) | -0.02 (-0.31 to 0.27) | 0.88 |
| | 9 months | 2.38 (1.25) | 2.43 (1.46) | -0.06 (-0.36 to 0.24) | 0.70 |
| | 12 months | 2.40 (1.50) | 2.42 (1.47) | -0.01 (-0.31 to 0.29) | 0.93 |
| **MPI Life control**** | Overall effect | | | 0.00 (-0.20 to 0.20) | 0.98 |
| | Baseline | 3.92 (1.10) | 4.18 (0.88) | | |
| | 3 months | 4.26 (1.06) | 4.17 (0.96) | 0.17 (-0.08 to 0.41) | 0.18 |
| | 6 months | 4.13 (1.11) | 4.32 (1.01) | -0.15 (-0.41 to 0.10) | 0.23 |
| | 9 months | 4.26 (1.07) | 4.31 (1.05) | -0.00 (-0.25 to 0.25) | 0.99 |
| | 12 months | 4.28 (1.03) | 4.32 (1.08) | -0.02 (-0.28 to 0.23) | 0.86 |

| | | Mean Intervention group (SD) | Mean Control group (SD) | Treatment effect (95%CI) | P value for difference |
|---|---|---|---|---|---|
| **MPI Affective distress**** | Overall effect | | | 0.03 (-0.10 to 0.16) | 0.69 |
| | Baseline | 2.71 (0.81) | 2.60 (0.74) | | |
| | 3 months | 2.56 (0.70) | 2.46 (0.73) | 0.07 (-0.11 to 0.24) | 0.46 |
| | 6 months | 2.55 (0.74) | 2.52 (0.77) | -0.03 (-0.20 to 0.15) | 0.78 |
| | 9 months | 2.49 (0.73) | 2.43 (0.68) | 0.06 (-0.12 to 0.24) | 0.51 |
| | 12 months | 2.47 (0.68) | 2.48 (0.64) | -0.00 (-0.19 to 0.18) | 0.97 |
| **MPI Support**** | Overall effect | | | 0.06 (-0.14 to 0.26) | 0.56 |
| | Baseline | 4.60 (0.97) | 4.42 (1.15) | | |
| | 3 months | 4.43 (1.06) | 4.35 (1.27) | -0.03 (-0.29 to 0.24) | 0.85 |
| | 6 months | 4.34 (1.24) | 4.22 (1.34) | 0.02 (-0.25 to 0.29) | 0.88 |
| | 9 months | 4.36 (1.29) | 4.17 (1.34) | 0.12 (-0.15 to 0.39) | 0.38 |
| | 12 months | 4.37 (1.20) | 4.15 (1.51) | 0.15 (-0.13 to 0.42) | 0.30 |
| **RAND-36 Physical health***** | Overall effect | | | -0.42 (-4.11 to 3.26) | 0.82 |
| | Baseline | 46.20 (19.18) | 47.20 (16.88) | | |
| | 3 months | 57.67 (21.23) | 53.79 (20.72) | 3.41 (-0.89 to 7.71) | 0.12 |
| | 6 months | 57.68 (20.71) | 56.85 (20.24) | 0.21 (-4.15 to 4.56) | 0.93 |
| | 9 months | 56.89 (21.84) | 58.70 (19.38) | -2.07 (-6.47 to 2.32) | 0.35 |
| | 12 months | 57.30 (22.99) | 58.87 (20.32) | -4.02 (-8.45 to 0.40) | 0.07 |
| **RAND-36 mental health***** | Overall effect | | | -0.69 (-3.35 to 1.96) | 0.61 |
| | Baseline | 73.68 (15.41) | 75.24 (14.82) | | |
| | 3 months | 75.52 (17.31) | 75.96 (15.77) | -0.66 (-3.91 to 2.60) | 0.69 |
| | 6 months | 77.36 (16.42) | 77.46 (15.66) | -0.18 (-3.49 to 3.14) | 0.92 |
| | 9 months | 76.75 (16.85) | 77.15 (16.17) | -1.68 (-5.03 to 1.67) | 0.33 |
| | 12 months | 77.98 (15.92) | 76.84 (16.01) | -0.26 (-3.64 to 3.13) | 0.88 |

*Values presented are model estimates of linear mixed-effects models with a random intercept, and adjusted for baseline and age, gender, BMI, education, smoking, marital status, back pain complaint history, patient expectations and baseline values. Regression coefficients can be interpreted as mean differences between interventions at a certain follow-up moment compared to baseline.*
*Abbreviations: EQ5D, Utility scores based on the EuroQol5D; MPI, Multidimensional Pain Inventory; RAND-36, Research and Development 36 item health survey.*
*** Higher score indicates more severe symptoms. Range for NRS pain, 0-10; for ODI, 0-100; for GPE, 1-7; for patient satisfaction, 1-7, for MPI 0-168.*
**** Higher score indicates more quality of life. Range for EQ5D utility: 0-1; for RAND36 0-100.*

**Cost-utility analyses**

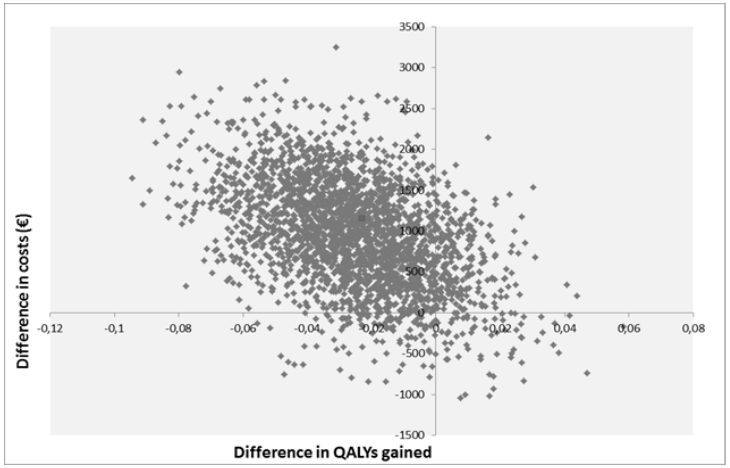
The RF denervation costs were estimated at €909 per participant. All other between-group differences in total and disaggregate costs were not statistically significant different (Table 5).

For QALYs, the ICER was -50036. This indicates that one QALY lost was associated with a societal cost of €50.036 (Table 6). More than 80% of the CE-pairs were located in the north-west quadrant, indicating that RF denervation was on average less effective and more costly. The CEAC (Figure 2) indicates that the maximum probability of the intervention being cost-effective was low (<0.1), irrespective of the willingness-to-pay. The overall conclusion of this study would not change when using any of the assumptions of the sensitivity analyses (Table 6).

***Table 5. Mean costs per participant in the intervention and control group, and mean cost differences between both groups during the 12-month follow-up period from a societal perspective***

| Cost category | Intervention group n=125; mean (SEM) | Control group n=126; mean (SEM) | Mean cost difference crude (95%CI) | Mean cost difference adjusted (95%CI) |
|---|---|---|---|---|
| **Intervention** | 909.00 (0.00) | NA | NA | NA |
| **Primary healthcare** | 1512.73 (143.65) | 1382.48 (110.07) | 130.25 (-210.37 to 507.18) | 90.34 (-269.45 to 503.06) |
| **Secondary healthcare** | 775.27 (114.73) | 844.85 (106.92) | -69.57 (-361.92 to 209.69) | -1.90 (-279.90 to 299.24) |
| **Medication** | 111.26 (20.44) | 93.08 (21.56) | 18.18 (-40.95 to 74.72) | 25.15 (-41.71 to 87.27) |
| **Informal care** | 778.03 (144.17) | 756.48 (182.40) | 21.54 (-451.29 to 462.52) | 91.45 (-344.93 to 531.73) |
| **Absenteeism** | 547.67 (202.66) | 579.20 (190.27) | -31.53 (-501.64 to 543.68) | -29.72 (-568.14 to 665.15) |
| **Unpaid productivity** | 1019.42 (188.89) | 1130.80 (201.83) | -111.38 (-648.50 to 430.82) | 65.61 (-484.53 to 644.64) |
| **Total** | **5653.38 (4730.26)** | **4786.89 (452.80)** | **866.50 (-358.89 to 2145.91)** | **1184.51 (-78.10 to 2472.40)** |

*Abbreviations: n, number; SEM, Standard Error of the Mean; CI, Confidence Interval; NA, Not Applicable; SD, Standard Deviation Note: Costs are expressed in 2014 Euros*



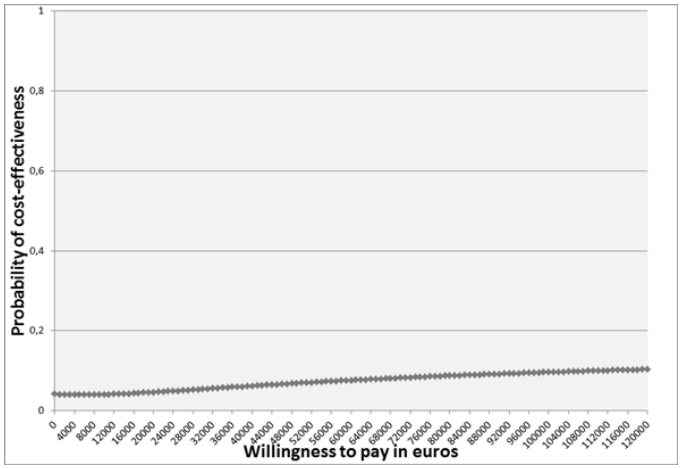




***Figure 2. Cost-effectiveness planes indicating the distribution of incremental cost-effect pairs around its four quadrants (1) and cost-effectiveness acceptability curves indicating the probability of adding radiofrequency denervation added to a standardised exercise programme being cost-effective in comparison with a standardised exercise programme alone for different values (€) of willingness-to-pay (2) for QALYS (EQ5D)***

***Table 6: Differences in pooled mean costs and effects (95% Confidence intervals), incremental cost-effectiveness ratios, and the distribution of incremental cost-effect pairs around the quadrants of the cost-effectiveness planes***

| Analysis | Sample size | | Outcome | ΔC (95% CI) | ΔE (95% CI) | ICER | Distribution CE-plane (%) | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Int. | Control | | € | Point | €/point | NE | SE | SW | NW |
| **Main analysis – Imputed dataset** | 125 | 126 | QALY EQ5D (0-1) | 1149.94 (-104.17 to 2476.29) | -0.02 (-0.07 to 0.02) | -50036 | 14.0 | 1.9 | 2.2 | 81.9 |
| **SA1 – complete cases** | 84 | 88 | QALY EQ5D (0-1) | 1025.01 (-314.02 to 2515.26) | -0.004 (-0.05 to 0.04) | -274175 | 39.1 | 5.4 | 1.4 | 54.1 |
| **SA2 - HCA** | 125 | 126 | QALY EQ5D (0-1) | 1149.94 (-104.17 to 2476.29) | -0.02 (-0.07 to 0.02) | -50035 | 14.0 | 1.9 | 2.2 | 81.9 |
| **SA3 – SF6D** | 125 | 126 | QALY SF6D (0-1) | 1149.94 (-104.17 to 2476.29) | -0.0008 (-0.01 to 0.01) | -1531177 | 43.6 | 2.6 | 1.5 | 52.4 |

*Abbreviations; CI: Confidence Interval; ICER: Incremental Cost-Effectiveness Ratio; CE-lane: Cost-Effectiveness plane; SA: Sensitivity Analysis; HCA: Human Capital Approach. Note: Costs are expressed in 2014 Euros*

## DISCUSSION

There were no statistically significant differences between RF denervation added to a standardised exercise programme compared to a standardised exercise programme alone, in patients with CLBP originating from the facet joints, for pain intensity, functional status, GPE, costs or any secondary outcome measure at our primary endpoint (three-month follow-up). Statistically significant results in favour of the intervention group were found only for success based on GPE at three weeks (OR 7.04; 95%CI: 2.46 to 20.12) and six weeks (3.28; 95%CI: 1.43 to 7.52) follow-up.

Strengths of this trial are the large sample and stratified randomisation that allowed for well-balanced study groups which can provide precise outcomes, the adequate one-year follow-up, and pragmatic design set in routine clinical care which is also considered the best design for an economic evaluation.

The non-blinded study design is a potential limitation of this trial that might have influenced the results. Evidence suggests that treatment effects may be exaggerated for subjective outcomes, when outcome assessors are not blinded.[39] However, this would most likely have led to an overestimation of the treatment effect and as the difference in effect between the two groups is rather small in this study, risk of bias due to the lack of blinding would not change the conclusion. Another possible limitation is the lack of a widely agreed gold standard for diagnosing facet joint pain. In this trial, 50% pain improvement after a single medial branch block was used as diagnostic criterion. Controversy about the ideal cut-off value in terms of selecting patients for RF denervation and the use of single or double diagnostic blocks exists. A 50% cut-off is mostly used in previous performed RCTs and in standard practice.[17,40] Performing multiple blocks will decrease the false-positive rate, but increase the number of false-negative blocks.[41] Furthermore, one trial performed in the USA, supports the notion that double-block tests are not effective or cost-effective.[42] Possible limitations for the economic evaluation are the assessment of all effect measures and cost-effect measures using retrospective questionnaires. This may have introduced recall bias. However, we tried to limit this by minimising the recall period to three months.[27] Furthermore, it is important to keep in mind that economic evaluation results are not transferable between countries due to differences in healthcare and/or social security systems.[43]

Recent systematic reviews have shown moderate quality evidence that facet joint RF denervation has a greater effect on pain compared with placebo. Low quality

evidence showed that facet joint RF denervation is more effective for pain than steroid injections.[11-13] In the trials included in this review, patients had a one point lower baseline pain score compared to our trial, but the RF denervation groups decreased more (to on average 3.3 out of 10 after three months) than the placebo group (to 5.0 out of 10 after three months). In our study, participants in both groups decreased in pain, but remained on a higher pain level compared to the RF denervation group in other studies in the systematic review. Kamper et al.[44] showed in a large systematic review that patients with CLBP receiving multidisciplinary rehabilitation experience less pain and disability than those receiving usual care or a physical treatment. In the trials in this review, overall pain scores were lower compared to our study as well. The population in our study had more severe baseline pain (although the same disability scores) compared to other trials. Furthermore, the participants in this study had very long-standing symptoms. As a longer duration of symptoms is an indicator of poor prognosis, this might be explanatory for the remaining pain. No pragmatic trials or comparisons on cost-effectiveness are available and the findings from this large pragmatic trial are therefore hard to compare with the current literature. The effectiveness of non-interventional treatments and other treatments options in patients with facet joint pain need to be evaluated. At the moment we have vast data included in the observational part of the MinT study. This is an opportunity for future research to identify subgroups of patients that might benefit from RF denervation or other treatments; and to examine diagnostic or prognostic tests for CLBP in a usual care setting.

This study shows that RF denervation for patients with CLBP originated from the facet joints does not have additional value to a standardised exercise programme compared to a standardised exercise programme alone, and is neither cost-effective. Although participants in both groups improved, RF denervation had no additional effect on pain, functional status, GPE, costs or any other secondary outcome.



**Acknowledgements**

The authors thank the Netherlands Organisation for Health Research and Development (171202013), the Dutch Society for Anaesthesiology, and the Dutch health insurance companies for funding this study. Merel van Raamt and Robert van Cingel (Sports Medical Centre Papendal) for their help in developing the exercise programme. The authors would also like to thank all participants, physiotherapists, and everyone working on the MinT study at one of the participating pain clinics.

## REFERENCES


1. Martin B, Deyo R, Mirza S, et al. Expenditures and health status among adults with back and neck problems. *JAMA*. 2008; 299:656-64
2. Manchikanti L, Singh V, Datta S, Cohen S, Hirsch J. Comprehensive review of epidemiology, scope, and impact of spinal pain. *Pain Physician.* 2009; 12:35-70
3. Hoy D, March L, Brooks P, et al. The global burden of low back pain: estimates from the Global Burden of Disease 2010 study. *Ann Rheum Dis.* 2014;73:968-974.
4. Vos T, Flaxman AD, Naghavi M, et al. Years lived with disability (YLDs) for 1160 sequelae of 289 diseases and injuries 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. The *Lancet.* 2013; 9859:2163-96
5. Lambeek L, Van Mechelen W, Knol D, Loisel P, Anema J. Randomised controlled trial of integrated care to reduce disability from chronic low back pain in working and private life. *BMJ*. 2010; 340:1050-1058
6. Goldthwait J. The Lumbo-Sacral Articulation; An Explanation of Many Cases of" Lumbago,"" Sciatica" and Paraplegia. *Boston Med Surg J.* 1911; 164:365-72
7. Ghormley R. Low back pain with special reference to the articular facets, with presentation of an operative procedure. *JAMA*. 1933; 101:1773-77
8. Cohen S, Raja S. Pathogenesis, diagnosis, and treatment of lumbar zygapophysial (facet) joint pain. *Anesthesiology.* 2007; 106:591-14
9. Cosman E. Electric and thermal field effects in tissue around radiofrequency electrodes. *Pain Med.* 2005; 6:405-24
10. Kline M. Radiofrequency techniques in clinical practice. Interventional Pain Management. In: Waldman S, Winnie A, eds. *Interventional Pain Mangement.* PhilidelphiaL Sanders; 1996: 1985-217.
11. Poetscher AW, Gentil AF, Lenza M, Ferretti M. Radiofrequency denervation for facet joint low back pain: a systematic review. *Spine*; 2014; 39:E842-E49
12. Center P, Manchikanti L. A Systematic Review and Best Evidence Synthesis of Effectiveness of Therapeutic Facet Joint Interventions in Managing Chronic Spinal Pain*. Pain Physician*. 2015; 18:E535-E82
13. Maas ET, Ostelo RWJG, Niemisto L, Jousimaa J, Hurri H, Malmivaara A, van Tulder MW. Radiofrequency denervation for chronic low back pain. *Cochrane Database Syst Rev* 2015, Issue 10. Art. No.: CD008572. DOI: 10.1002/14651858.CD008572.pub2.
14. Maas E, Juch J, Groeneweg J, Ostelo R,et al. Cost-effectiveness of minimal interventional procedures for chronic mechanical low back pain: design of four randomised controlled trials with an economic evaluation. *BMC Musculoskelet Disord.* 2012. 13:260
15. Ostelo R, Deyo R, Stratford P, Waddell G, Croft P, Von Korff M, Bouter L, Henrica C. Interpreting change scores for pain and functional status in low back pain: towards international consensus regarding minimal important change. *Spine.* 2008; 33:90-94
16. Bekkering G, Hendriks H, Koes B, et al. KNGF Guideline low back Pain. *Ned Tijdschr Fysiother.* 2005; 111:1-24.
17. Itz C, Willems P, Zeilstra D, Huygen F. Dutch Multidisciplinary Guideline for Invasive Treatment of Pain Syndromes of the Lumbosacral Spine. *Pain Practice.* 2016; 16(1):90-110.

18. Downie W, Leatham P, Rhind V, Wright V, Branco J, Anderson J. Studies with pain rating scales. *Ann Rheum Dis*. 1978; 37:378-81
19. Devilly G, Borkovec T. Psychometric properties of the credibility/expectancy questionnaire. *J Behav Ther Exp Psychiatry*. 2000; 31:73-86
20. Fairbank J, Pynsent P. The Oswestry disability index. *Spine.* 2000; 25:2940-53
21. Chiarotto A, Deyo RA, Terwee CB, et al.. Core outcome domains for clinical trials in non-specific low back pain. *Eur Spine J.* 2015; 24:1127-42
22. Kind P. The EuroQoL instrument: an index of health-related quality of life. *Quality of life and pharmacoeconomics in clinical trials*. Edited by Spilker B. Philadelphia: Lippincott-Raven Publihsers; 1996:191-201
23. Bombardier C. Outcome assessments in the evaluation of treatment of spinal disorders: summary and general recommendations. *Spine.* 2000; 25:3100-03
24. Hays R, Morales L. The RAND-36 measure of health-related quality of life. *Ann Med*. 2001; 33:350-57
25. Lousberg R, Van Breukelen G, Groenman N, Schmidt A, Arntz A, Winter F. Psychometric properties of the Multidimensional Pain Inventory, Dutch language version (MPI-DLV). *Behav Res Ther* 1999; 37:167-82
26. Lamers LM SP, McDonnell J, Krabbe PFM, van Busschbach JJ. Kwaliteit van leven met in economische evaluaties: het Nederlands EQ-5D tarief. *Ned Tijdschr Geneeskd.* 2005; 149:1574-78
27. Goossens M, Rutten-van Mölken M, Vlaeyen J, van der Linden S. The cost diary: a method to measure direct and indirect costs in cost-effectiveness research. *J Clin Epidemiol.* 2000; 53:688-95
28. Hakkaart-van Roijen L, Tan S, Bouwmans C. Handleiding voor kostenonderzoek: Methoden en standaard kostprijzen voor economische evaluaties in de gezondheidszorg. College voor zorgverzekeringen; 2011.
29. Riegelman RK. Studying a study and testing a test: how to read the medical evidence. Lippincott Williams & Wilkins, Philadelphia; 2005
30. Koopmanschap MA. PRODISQ: a modular questionnaire on productivity and disease for economic evaluation studies. *Expert Rev Pharmacoecon Outcomes Res*. 2005; 5:23-28
31. Manchikanti L, Pampati V, Fellows B, Bakhit CE. The diagnostic validity and therapeutic value of lumbar facet joint nerve blocks with or without adjuvant agents. *Curr Rev Pain.* 2000; 4:337-44
32. Twisk J. Applied multilevel analysis: a practical guide for medical researchers. Cambridge Univ Press, Cambridge; 2006
33. Rubin D. Multiple imputation for nonresponse in surveys, vol 81. John Wiley & Sons, Hoboken New Jersey, 2004
34. Groenwold R, Donders AR, Roes K, Harrell F, Moons K. Dealing with missing outcome data in randomized trials and observational studies. *Am J Epidemiol*. 2012; 175: 210-7
35. Sterne JA, White IR, Carlin JB, et al. Multiple imputation for missing data in epidemiological and clinical research: potential and pitfalls. *BMJ*. 2009. 338:b2393
36. Willan A, Briggs A, Hoch J. Regression methods for covariate adjustment and subgroup analysis for non-censored cost-effectiveness data. *Health economics*. 2004. 13:461-75
37. Black W. The CE plane a graphic representation of cost-effectiveness. *Med Decis Making*. 1990. 10:212-14

38. Fenwick E, O'Brien B, Briggs A. Cost-effectiveness acceptability curves–facts, fallacies and frequently asked questions. *Health economics*. 2004. 13:405-15
39. Kahan BC, Cro S, Doré CJ, et al. Reducing bias in open-label trials where blinded outcome assessment is not feasible: strategies from two randomised trials. *Trials*. 2014. 15:456
40. Cohen SP, Huang JH, Brummett C. Facet joint pain—advances in patient selection and treatment. *Nat Rev Rheum*. 2013. 9:101-16
41. Boswell MV, Singh V, Staats PS, Hirsch JA. Accuracy of precision diagnostic blocks in the diagnosis of chronic spinal pain of facet or zygapophysial joint origin: A systematic review. *Pain Physician*. 2013. 6:449-56
42. Bogduk N, Holmes S. Controlled Zygapophysial Joint Blocks: The Travesty of Cost-Effectiveness. *Pain Med*. 2000. 1:24-34
43. Tompa E, Culyer AJ, Dolinschi R. Economic evaluation of interventions for occupational health and safety: developing good practice. Oxford University Press, Oxford, 2008
44. Kamper SJ, Apeldoorn A, Chiarotto A, et al. Multidisciplinary biopsychosocial rehabilitation for chronic low back pain: Cochrane systematic review and meta-analysis. *BMJ;* 2015; 350:h444

***Table S1. Baseline characteristics of completers versus non-completers***

| Characteristics | Intervention-Randomised: N=125 Complete baseline: N=117 | Intervention Complete N=78 Complete baseline: N=78 | Intervention Incomplete N=47 Complete baseline:\| N=39 | Control Randomised: N=126 Complete baseline: N=116 | Control Complete N=88 Complete baseline: N=88 | Control Incomplete N=38 Complete baseline: N=28 |
|---|---|---|---|---|---|---|
| **Age in years (SD)** | 52.9 (11.5) | 54.4 (10.9) | 50.1 (12.0) | 52.6 (10.8) | 53.2 (10.5) | 50.7 (11.8) |
| **Female (N (%))** | 65 (55.6%) | 45 (57.7%) | 19 (48.7%) | 60 (51.7%) | 48 (54.5%) | 16 (57.1%) |
| **BMI (SD)** | 26.7 (5.2) | 27.2 (5.7) | 25.9 (3.8) | 27.6 (4.3) | 27.9 (4.3) | 26.6 (4.1) |
| **Smoker (N (%))** | 34 (29.1%) | 16 (20.5%) | 18 (46.2%) | 34 (29.3%) | 22 (25.0%) | 12 (42.9%) |
| **Education** | | | | | | |
| • Low (N (%)) | 57 (48.7%) | 40 (51.3%) | 18 (48.6%) | 64 (55.2%) | 48 (54.5%) | 17 (60.7%) |
| • Moderate (N (%)) | 35 (29.9%) | 22 (28.2%) | 13 (35.1%) | 34 (29.3%) | 27 (30.7%) | 8 (28.6%) |
| • High (N (%)) | 21 (17.9%) | 16 (20.5%) | 6 (16.2%) | 16 (13.8%) | 13 (14.8%) | 3 (10.7%) |
| **History of back pain complaints** | | | | | | |
| • Time since first experience with low back pain in months (median (IQR)) | 146 (50-267) | 158 (43 – 304) | 146 (55-220) | 100 (37-186) | 115 (36-186) | 84 (50 – 220) |
| • Time since current episode with low back pain in months (median (IQR)) | 31 (12-103) | 30 (12 – 78) | 37 (8-131) | 27 (11-73) | 29 (12 – 86) | 20 (7-70) |
| **Married/living with a partner (N (%))** | 93 (74.4%) | 68 (87.2%) | 25 (64.1%) | 98 (77.8%) | 76 (86.4%) | 22 (78.6%) |
| **Expectations** | | | | | | |
| • Credibility (0-27) | 21.4 (3.9) | 21.8 (3.4) | 20.4 (4.8) | 19.5 (5.5) | 19.2 (5.9) | 20.3 (4.1) |
| • Expectancy (0-27) | 18.9 (4.6) | 19.4 (4.3) | 18.2 (5.0) | 17.4 (5.2) | 16.8 (5.6) | 19.0 (3.2) |
| **Having a paid job** | 64 (51.2%) | 40 (55.6%) | 24 (53.3%) | 66 (58.9%) | 52 (59.1%) | 14 (58.3%) |
| **Outcomes** | | | | | | |
| **Pain intensity in the past week (NRS 0-10)** | 7.1 (1.4) | 6.9 (1.5) | 7.4 (1.1) | 7.2 (1.3) | 7.1 (1.3) | 7.4 (1.4) |
| **Oswestry disability index (mean (SD))** | 35.1 (14.7) | 35.0 (14.0) | 35.2 (16.1) | 34.4 (12.2) | 33.9 (11.5) | 36.0 (14.4) |
| **Quality of life (EQ-5D)** | 0.5 (0.3) | 0.5 (0.2) | 0.5 (0.3) | 0.5 (0.3) | 0.5 (0.3) | 0.5 (0.3) |

*Abbreviations: SD, Standard Deviation; IQE, Inter Quartile Range; N, number; BMI, Body Mass Index; EQ-5D, EuroQol-5D*

***Table S2. As treated analysis for primary outcomes, without 12 protocol violators based on intention-to-treat analyses***

| | | Mean Intervention group (N=125) (SD) | Mean Control group (N=114) (SD) | Treatment effect (95%CI) | P value for difference | |
|---|---|---|---|---|---|---|
| **Primary outcomes** | | | | | | |
| **NRS Pain*** | Overall effect | | | -0.21 (-0.62 to 0.20) | 1.69 | |
| | Baseline | 7.14 (1.38) | 7.22 (1.31) | | | |
| | 3 weeks | 5.17 (2.27) | 5.99 (1.69) | -0.53 (-1.13 to 0.08) | 1.91 | |
| | 6 weeks | 5.19 (2.31) | 5.95 (1.78) | -0.50 (-1.08 to 0.08) | 1.91 | |
| | 3 months | 5.01 (2.29) | 5.50 (2.13) | -0.29 (-0.87 to 0.29) | 1.68 | |
| | 6 months | 4.61 (2.89) | 4.92 (2.32) | -0.12 (-0.72 to 0.47) | 1.32 | |
| | 9 months | 4.66 (2.41) | 4.95 (2.43) | -0.05 (-0.65 to 0.55) | 1.13 | |
| | 12 months | 4.49 (2.48) | 4.56 (2.51) | 0.31 (-0.29 to 0.92) | 0.31 | |
| **ODI Functioning*** | Overall effect | | | 0.09 (-3.76 to 3.93) | 0.96 | |
| | Baseline | 35.08 (14.66) | 34.74 (11.85) | | | |
| | 3 months | 26.03 (16.58) | 29.61 (15.26) | -2.29 (-6.52 to 1.93) | 1.71 | |
| | 6 months | 25.38 (15.64) | 27.75 (15.97) | -0.58 (-4.84 to 3.68) | 1.21 | |
| | 9 months | 25.74 (15.55) | 24.97 (15.56) | 2.23 (-2.05 to 6.52) | 0.31 | |
| | 12months | 24.59 (16.28) | 25.62 (16.80) | 1.47 (-2.84 to 5.77) | 0.50 | |
| | | **Success % Intervention group** | **Success % Control group** | **OR (95%CI)** | **P value for difference** | **NNT** |
| **GPE Success** | Overall effect | | | 1.69 (1.12 to 2.57) | 0.01 | |
| | 3 weeks | 29.6 | 3.3 | 11.82 (3.33 to 42.02) | 0.0001 | 4 |
| | 6 weeks | 29.4 | 6.5 | 5.47 (2.15 to 13.97) | 0.0004 | 5 |
| | 3 months | 36.1 | 22.1 | 1.86 (0.91 to 3.79) | 0.09 | 7 |
| | 6 months | 40.7 | 33.3 | 1.30 (0.65 to 2.58) | 0.46 | 14 |
| | 9 months | 38.7 | 37.5 | 0.87 (0.43 to 0.74) | 0.69 | 84 |
| | 12months | 42.7 | 37.9 | 0.98 (0.49 to 1.97) | 0.96 | 21 |

*Values presented are model estimates of linear mixed-effects models with a random intercept, and adjusted for baseline and age, gender, BMI, education, smoking, marital status, back pain complaint history, patient expectations and baseline values. Regression coefficients can be interpreted as mean differences between interventions at a certain follow-up moment compared to baseline.Abbreviation: SD: Standard Deviation; NRS, Numeric Rating Scale; GPE, Global Perceived Effect; NNT: Number Needed to Treat*
** Higher score indicates more severe symptoms. Range for NRS pain, 0-10; for GPE, 1-7; ODI, 0-100*

***Table S3. Treatment effects for primary outcomes based on an as-treated after 3 months***

| | | Mean Intervention group (SD) | Mean Control group (SD) | Treatment effect (95%CI) | P value for difference | |
|---|---|---|---|---|---|---|
| **Primary outcomes** | | | | | | |
| **NRS Pain*** | Overall effect | | | 0.10 (-0.39 to 0.58) | 0.70 | |
| | Baseline | 7.14 (1.38) | 7.07 (1.31) | | | |
| | 3 weeks | 5.17 (2.27) | 5.60 (1.83 | -0.13 (-0.83 to 0.57) | 1.28 | |
| | 6 weeks | 5.19 (2.31) | 5.60 (2.19) | -0.14 (-0.8 to 0.52) | 1.32 | |
| | 3 months | 5.01 (2.29) | 5.23 (2.31) | -0.02 (-0.68 to 0.64) | 1.05 | |
| | 6 months | 4.61 (2.89) | 4.62 (2.37) | 0.16 (-0.52 to 0.84) | 0.64 | |
| | 9 months | 4.66 (2.41) | 4.52 (2.49) | 0.28 (-0.41 to 0.97) | 0.42 | |
| | 12 months | 4.49 (2.48) | 4.28 (2.33) | 0.55 (-0.15 to 1.2) | 0.12 | |
| **ODI Functioning*** | Overall effect | | | 0.89 (-2.79 to 4.58) | 0.64 | |
| | Baseline | 35.08 (14.66) | 33.89 (13.35) | | | |
| | 3 months | 26.03 (16.58) | 27.77 (16.74) | -0.86 (-1.27 to -0.45) | 2.00 | |
| | 6 months | 25.38 (15.64) | 26.45 (16.99) | 0.30 (-3.87 to 4.47) | 0.89 | |
| | 9 months | 25.74 (15.55) | 23.84 (16.88) | 2.21 (-2.00 to 6.42) | 0.30 | |
| | 12months | 24.59 (16.28) | 23.77 (16.27) | 2.42 (-1.81 to 6.65) | 0.26 | |
| | | **Success % Intervention group** | **Success % Control group** | **OR (95%CI)** | **P value for difference** | **NNT** |
| **GPE Success** | Overall effect | | | 1.09 (0.69 to 1.73) | 0.71 | |
| | 3 weeks | 29.6 | 8.6 | 3.81 (1.29 to 11.29) | 0.02 | 5 |
| | 6 weeks | 29.4 | 11.1 | 2.72 (1.08 to 6.87) | 0.03 | 6 |
| | 3 months | 36.1 | 31.0 | 1.03 (0.48 to 2.20) | 0.94 | 20 |
| | 6 months | 40.7 | 39.4 | 0.85 (0.40 to 1.82) | 0.68 | 77 |
| | 9 months | 38.7 | 43.5 | 0.63 (0.29 to 1.38) | 0.25 | -21 |
| | 12 months | 42.7 | 45.9 | 0.63 (0.29 to 1.38) | 0.25 | -32 |

*Values presented are model estimates of linear mixed-effects models with a random intercept. and adjusted for baseline and age. gender. BMI. education. smoking. marital status. back pain complaint history. patient expectations and baseline values. Regression coefficients can be interpreted as mean differences between interventions at a certain follow-up moment compared to baseline.*
*Abbreviation: SD: Standard Deviation NRS. Numeric Rating Scale; GPE. Global Perceived Effect; NNT: Number Needed to Treat*
** Higher score indicates more severe symptoms. Range for NRS pain. 0-10; for GPE. 1-7; ODI. 0-100*

***Table S4. Treatment effects for complete cases for primary outcomes based on intention-to-treat analyses***

| | | Mean Intervention group (SD) | Mean Control group (SD) | Treatment effect (95%CI) | P value for difference | |
|---|---|---|---|---|---|---|
| **Primary outcomes** | | | | | | |
| **NRS Pain*** | Overall effect | | | -0.44 (-0.96 to 0.07) | 0.09 | |
| | Baseline | 6.94 (1.51) | 7.14 (1.27) | | | |
| | 3 weeks | 5.15 (2.24) | 5.86 (1.81) | -0.54 (-1.25 to 0.18) | 0.14 | |
| | 6 weeks | 4.94 (2.23) | 5.89 (1.96) | -0.77 (-1.48 to -0.06) | 0.03 | |
| | 3 months | 4.47 (2.13) | 5.21 (2.28) | -0.57 (-1.28 to 0.15) | 0.12 | |
| | 6 months | 4.32 (2.25) | 4.81 (2.42) | -0.32 (-1.03 to 0.39) | 0.38 | |
| | 9 months | 4.17 (2.31) | 4.84 (2.49) | -0.50 (-1.21 to 0.21) | 0.17 | |
| | 12months | 4.18 (2.51) | 4.31 (2.54) | 0.04 (-0.67 to 0.75) | 0.91 | |
| **ODI Functioning*** | Overall effect | | | -1.17 (-4.89 to 2.55) | 0.54 | |
| | Baseline | 34.28 (13.86) | 33.47 (11.16) | | | |
| | 3 months | 23.31 (14.93) | 26.90 (15.26) | -3.60 (-7.83 to 0.63) | 0.10 | |
| | 6 months | 24.06 (15.23) | 25.10 (15.05) | -1.05 (-5.28 to 3.18) | 0.63 | |
| | 9 months | 23.58 (14.28) | 24.25 (14.36) | -0.67 (-4.90 to 2.56) | 0.76 | |
| | 12months | 24.22 (16.01) | 23.57 (15.06) | 0.64 (-3.59 to 4.87) | 0.77 | |
| | | **Success % Intervention group** | **Success % Control group** | **OR (95%CI)** | **P value for difference** | **NNT** |
| **GPE Success** | Overall effect | | | 1.67 (1.03 to 2.72) | 0.04 | |
| | 3 weeks | 30.6 | 6.3 | 6.51 (2.12 to 19.98) | 0.0011 | 4 |
| | 6 weeks | 27.8 | 10.0 | 3.38 (0.70 to 16.38) | 0.13 | 6 |
| | 3 months | 43.1 | 26.3 | 2.10 (0.93 to 2.75) | 0.07 | 6 |
| | 6 months | 43.1 | 37.5 | 1.24 (0.56 to 2.69) | 0.61 | 18 |
| | 9 months | 41.7 | 37.5 | 1.15 (0.52 to 2.53) | 0.72 | 24 |
| | 12months | 43.1 | 41.3 | 1.04 (0.76 2.27) | 0.92 | 56 |

*Values presented are model estimates of linear mixed-effects models with a random intercept. and adjusted for baseline and age. gender. BMI. education. smoking. marital status. back pain complaint history. patient expectations and baseline values. Regression coefficients can be interpreted as mean differences between interventions at a certain follow-up moment compared to baseline.Abbreviation: SD: Standard Deviation; NRS. Numeric Rating Scale; GPE. Global Perceived Effect; ODI. Oswestry Disability Index*

** Higher score indicates more severe symptoms. Range for NRS pain. 0-10; for ODI. 0-100; for GPE. 1-7*